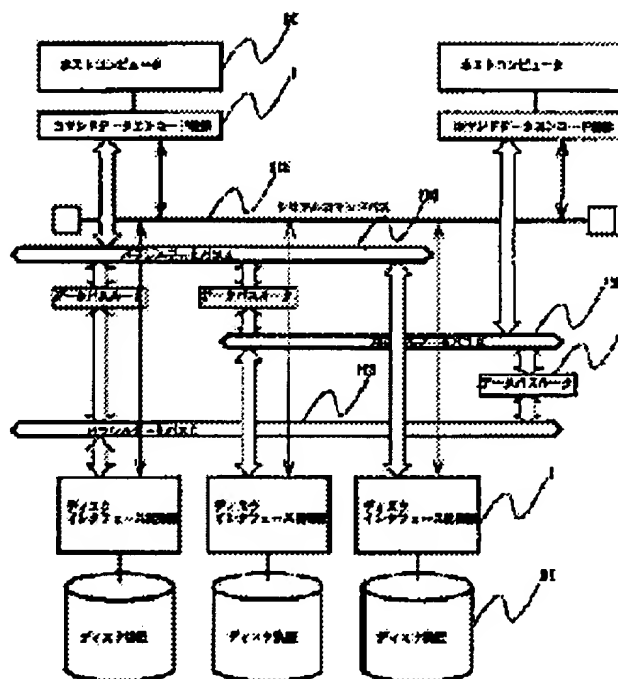


Patent number:	JP7191930
Publication date:	1995-07-28
Inventor:	SUGIMOTO KINICHI
Applicant:	NEC CORP
Classification:	
- international:	G06F13/10; G06F13/36
- european:	
Application number:	JP19930331351 19931227
Priority number(s):	

PURPOSE: To improve the throughput of a disk and the response of sequential data access by providing a bus for command processing and a bus for data processing independently of each other in the interface for disk device of a computer.

CONSTITUTION:This interface for disk device DK consists of a command data encoding mechanism 1 which generates a command packet from a corresponding disk device code and a disk I/O control command in response to a data read or write request from a host, a disk interface control mechanism 2 which receives the command packet issued from the command data encoding mechanism 1, a disk deice DK executed by the disk interface control mechanism, a route setting mechanism which discriminates the shortest transfer path, and a data path router 3 which continuously secures a path for the purpose of continuously transferring data.



<http://v3.espacenet.com/textdoc?DB=EPODOC&IDX=JP7191930>

Reference C

Japanese Patent Application Public-disclosure No. 7-191930

Japanese Patent Application Public-disclosure date: July 28, 1995

Title of the invention: Interface for a disk unit and a method of controlling the same

Japanese Patent Application No. 5-331351

Japanese Patent Application date: December 27, 1993

[Problems to be solved by the invention]

If bus overhead increases when accessing a large-capacity disk unit intensively managed by equipment such as a file server or the like, it is detrimental to processing performance of an entire system. In consideration of the problem, it is an object of the present invention to reduce overhead in transferring data, especially during processing not involving data transfer. It is a further object of the present invention, when a bus is occupied for data transfer, to enable a simultaneous operation of buses by handling precedences for using a data bus and utilizing multiple data transfer paths, to thereby reduce latency at the time of file access and improve response characteristics.

[Means for solving the problems]

An interface for disk units of computers and method of controlling the same in accordance with the present invention is characterized by comprising: a command data encode mechanism for generating a command to a disk unit in response to a data readout request/write request from a host computer; a serial command bus for transferring the command to the disk unit; a disk interface control means for receiving the command, separating an I/O control command of a disk from the received command and performing an input/output process to/from the disk unit; and a parallel data bus capable of transferring input/output data to/from the command data encode mechanism, independently of an operation of the command bus.

Further, an interface for disk units of computers and method of controlling the same in accordance with the present invention is characterized by comprising: a command data encode mechanism for generating a command packet from a corresponding disk unit code and I/O control command of a disk in response to a data readout/write request from a host computer; a disk interface control mechanism for receiving the command packet issued by the command data encode mechanism, the disk unit to be implemented by the disk interface control mechanism; a path setting mechanism for determining the shortest transfer path; and a data bus router for continuously securing a bus in order to continuously transfer data.

Further, a method of controlling disk units of computers in accordance with the present invention is characterized by securing a data bus recursively in order to transfer data via multiple data bus routers.

Still further, an interface for disk units of computers and method of controlling the same in accordance with the present invention is characterized by comprising: a data presence/absence determination mechanism for determining, in response to an input/output request transferred from a host computer to a disk interface control means, whether or not data exists on a disk unit coupled directly to the disk interface control means; a disk interface for requesting input to/output from the disk unit when it transpires that data exists on the disk unit; and an external interface access mechanism for issuing the input/output request from the host computer to another disk interface control means having the same function as the disk interface control means when it transpires that data does not exist on the disk unit.

[Function]

According to a conventional disk unit, an access to the disk unit not involving transfer of an actual data region such as an access for checking the status, reading a disk parameter or the like involves regular data transfer and thus requires a process such as bus arbitration or the like which is similar to a file access request. Since a conventional disk unit employs the same data transfer channel for the aforementioned process, it is difficult to improve the data transfer efficiency of a bus. According to a method of the present invention, a communication line is used on the basis of message passing when processing does not involve regular data transfer or short data is transferred, whereas a parallel interface used for data transfer only is secured when long data is transferred. By this method, according to the present invention, a data transfer operation and other operations can be performed in an overlapping manner, which improves not only bus use efficiency but also bus transfer speed.

Further, when multiple interfaces used exclusively for data transfer are secured, data transfer can be conducted by using another bus even if a particular disk unit is occupying a bus for a long time. Thus, at the time of a file access operation requiring sequential access such as display of digital moving images or the like, the other file access operations can be performed without any trouble, which provides disk units with satisfactory response characteristics. Similarly, it is possible to constitute a disk unit that does not add constraints to a system performance even in an environment where high-speed disk units and low-speed disk units coexist.

Still further, when multiple disk units are coupled to a single channel according

to a conventional interface for disk units, if a particular disk unit accesses the channel, the other disk units can no longer exchange data with other devices. According to a disk unit of the present invention, a disk unit is identified via a router, and then a bus is secured. Therefore, if a particular channel becomes inoperable to transfer data, a substitution path and disk unit can be found. Thus, the present invention can also improve fault tolerance.

[Brief explanation of the drawings]

Fig. 1 is a block diagram describing an example of a method of data transfer by conventional disk units.

Fig. 2 is a block diagram describing an embodiment of an automatic routing disk unit according to Claim 2 of the subject application.

Fig. 3 is a detail drawing of the disk interface control means described in Fig. 2.

Fig. 4 is a block diagram describing an embodiment of an automatic routing disk unit according to Claim 3 of the subject application.

Fig. 5 is a detail drawing of the data bus router described in Fig. 4.

Fig. 6 describes a format of a command packet used for executing a data transfer sequence defined in Claim 1 of the subject application.

Fig. 7 is a block diagram indicating a data transfer process sequence defined in Claim 1 of the subject application.

Fig. 8 is a block diagram describing an embodiment of an automatic routing disk unit in accordance with Claim 4 of the subject application.

Fig. 9 is a detail drawing of the disk interface control means described in Fig. 8.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-191930

(43) 公開日 平成7年(1995)7月28日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 13/10	3 4 0	8327-5B		
13/36	5 3 0 B	8944-5B		

審査請求 有 請求項の数 4 O L (全 10 頁)

(21) 出願番号 特願平5-331351

(22) 出願日 平成5年(1993)12月27日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 杉本 欽一

東京都港区芝五丁目7番1号 日本電気株式会社内

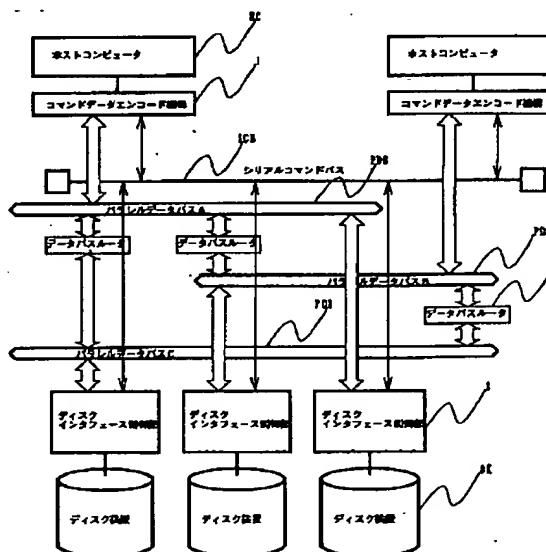
(74) 代理人 弁理士 京本 直樹 (外2名)

(54) 【発明の名称】 ディスク装置のインタフェース及びその制御方法

(57) 【要約】

【目的】 コンピュータのディスク装置用インタフェースのコマンド処理用バス、データ処理用バスを独立されることにより、ディスクのスループットの向上とシーケンシャルデータアクセスのレスポンス向上を目的とする。

【構成】 ディスク装置において、ホストからのデータの読みだし及び書き込み要求に対して、対応するディスク装置コードとディスクのI/O制御コマンドよりコマンドバケットを生成するコマンドデータエンコード機構と、前記コマンドデータエンコード機構より発行された前記コマンドバケットを受け取るディスクインタフェース制御機構と、前記ディスクインタフェース制御機構で実行される前記ディスク装置と、最短転送バスを判断する経路設定機構と、連続的にデータ転送をおこなうために連続的にバスを確保するデータバスルータからなるディスク装置のインタフェース。



1

【特許請求の範囲】

【請求項1】 コンピュータのディスク装置において、ホストコンピュータからのデータの読みだし及び書き込み要求に対して、ディスク装置へのコマンドを生成するコマンドデータエンコード機構と、前記コマンドを前記ディスク装置に対して転送するシリアルコマンドバスと、前記コマンドを引き取りディスクのI/O制御コマンドを分離して前記ディスク装置の入出力処理を実行するディスクインタフェース制御部と、入出力したデータを前記コマンドデータエンコード機構との間で前記コマンドバスの動作と独立してデータ転送可能なパラレルデータバスとからなるディスク装置のインタフェース。

【請求項2】 請求項1に記載のディスク装置において、ホストからのデータの読みだし及び書き込み要求に対して、対応するディスク装置コードとディスクのI/O制御コマンドよりコマンドバケットを生成するコマンドデータエンコード機構と、前記コマンドデータエンコード機構より発行された前記コマンドバケットを受け取るディスクインタフェース制御機構と、前記ディスクインタフェース制御機構で実行される前記ディスク装置と、最短転送バスを判断する経路設定機構と、連続的にデータ転送をおこなうために連続的にバスを確保するデータバスルータからなるディスク装置のインタフェース。

【請求項3】 請求項2に記載のディスク装置の制御方法において、複数のデータバスルータを経由してデータ転送をおこなうため、再起的にデータバスの確保を行うことを特徴とする。ディスク装置のインタフェース制御方法。

【請求項4】 コンピュータのディスク装置において、ホストコンピュータからディスクインタフェース制御部に転送された入出力要求にたいして、前記ディスクインタフェース制御部に直接接続されたディスク装置上にデータ存在するかを判断するデータ有無判定機構とデータがあった場合に前記ディスク装置に対して入出力を要求するディスクインタフェースと、前記ディスク装置上にデータが無かった場合に前記ディスクインタフェース制御部と同等の機能を持った他のディスクインタフェース制御部に対して、前記ホストコンピュータからの前記入出力要求を発行する外部インタフェースアクセス機構とからなるディスク装置のインタフェース。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明はコンピュータのディスク装置にかかわり、特に複数のディスク装置を使用して大量のデータ管理を行うディスク装置に関する。

【0002】

【従来の技術】 コンピュータシステムの扱うデータの大容量化が進み、ファイルサーバーなどの機器が管理するディスク装置の大容量化が著しい。従来はこのような環

2

境においても図1に示すような、ホストコンピュータと周辺装置とのインタフェースは単一のチャネルのI/Oバス（例えば、ANSI X3/131-1986 "Small Computer System Interface"）などで接続する場合が一般的であった。すなわち、各ホストコンピュータは最低限のI/Oバスインタフェースを経由してディスク装置などの外部記憶装置と接続される。その際に、ホストコンピュータからの外部記憶装置に対するアクセスは、I/Oバスのアービトレーション動作を含んだバスの占有を伴うため、データの転送を伴わないディスク装置のステータスの読みだしやディスク装置のパラメータの読み出しに関してもコマンド解釈などのソフトウェアオーバーヘッドやバスのアービトレーションに伴うハードウェアオーバーヘッドを伴うことになる。このようなデータの入出力を伴わないファイルアクセスの場合においてもI/Oバスを占有されるため他のデバイスあるいはホストコンピュータはデバイスに対するアクセスも不可能となる。そのため、オペレーティングシステムなどのファイル管理機能により、I/Oバスに対するアクセス回数を減らす管理を行うなどして、システムの向上を図るのが一般的であった。

【0003】

【発明が解決しようとしている課題】 ファイルサーバーなどの機器が集中的に管理する大容量のディスク装置に対するアクセスにおいて、特にバスのオーバーヘッドの割合が増加すると、システム全体の処理性能に支障をきたす。よって、データ転送実行時におけるオーバーヘッド、特にデータの転送を伴わない処理におけるオーバーヘッドの低減を目的とする。同時に、データ転送に伴うバス占有時に関しても、データバスの利用優先度の処理及び複数のデータ転送経路の使用により、バスの同時動作を可能にし、ファイルアクセス時のレイテンシの低減による応答性の向上を目的とする。

【0004】

【課題を解決するための手段】 本発明のディスク装置のインタフェース及びその制御方法は、コンピュータのディスク装置において、ホストコンピュータからのデータの読みだし及び書き込み要求に対して、ディスク装置へのコマンドを生成するコマンドデータエンコード機構と、前記コマンドを前記ディスク装置に対して転送するシリアルコマンドバスと、前記コマンドを引き取りディスクのI/O制御コマンドを分離して前記ディスク装置の入出力処理を実行するディスクインタフェース制御部と、入出力したデータを前記コマンドデータエンコード機構との間で前記コマンドバスの動作と独立してデータ転送可能なパラレルデータバスとからなることを特徴とする。

【0005】 また、本発明はコンピュータのディスク装置において、ホストからのデータの読みだし及び書き込

3

み要求に対して、対応するディスク装置コードとディスクのI/O制御コマンドよりコマンドバケットを生成するコマンドデータエンコード機構と、前記コマンドデータエンコード機構より発行された前記コマンドバケットを受け取るディスクインタフェース制御機構と、前記ディスクインタフェース制御機構で実行される前記ディスク装置と、最短転送バスを判断する経路設定機構と、連続的にデータ転送をおこなうために連続的にバスを確保するデータバスルータからなることを特徴とする。

【0006】また、本発明はコンピュータのディスク装置の制御方法において、複数のデータバスルータを経由してデータ転送をおこなうため、再起的にデータバスの確保を行うことを特徴とする。

【0007】また、本発明はコンピュータのディスク装置において、ホストコンピュータからディスクインタフェース制御部に転送された入出力要求にたいして、前記ディスクインタフェース制御部に直接接続されたディスク装置上にデータ存在するかを判断するデータ有無判定機構とデータがあった場合に前記ディスク装置に対して入出力を要求するディスクインタフェースと、前記ディスク装置上にデータが無かった場合に前記ディスクインタフェース制御部と同等の機能を持った他のディスクインタフェース制御部に対して、前記ホストコンピュータからの前記入出力要求を発行する外部インタフェースアクセス機構とからなることを特徴とする。

【0008】

【作用】従来ディスク装置に対するアクセスの内、ステータスのチェックやディスクパラメータの読みだし等の実際のデータ領域の転送を伴わないアクセスは、通常のデータ転送を伴いファイルアクセス要求と同様のバスアービトラージなどの処理を必要とする。そのため、通常は同じデータ転送チャンネルを使用していたが、結果としてバスのデータ転送効率の向上は困難であった。それに対し、本方法を適用することにより、通常データ転送を伴わない処理あるいは短い長さのデータ転送に対しては、メッセージパッシングを基本とした通信回線を使用し、長いデータ転送を伴うデータ転送に関してはデータ転送専用のパラレルインタフェースを確保することにより、データ転送動作とそれ以外の動作との処理のオーバーラップが可能となり、バスの使用効率の向上とバスの転送速度の向上が図られる。

【0009】また、データ転送専用インタフェースを複数確保した場合には、特定のディスク装置が長時間バスを占有している場合においても、他のバスを使用してのデータ転送が可能であるため、デジタル動画のようなシーケンシャルアクセスを必要とするファイルアクセス動作時においても、他のファイルアクセスに対する障害を生じないため、応答性のよいディスク装置を構成することが可能となる。同様に、高速のディスク装置と低速のディスク装置が混在するような環境においても、シス

4

テム性能を制約しない形態を構成可能となる。

【0010】また、従来のディスク装置用のインタフェースではチャンネル一つに対してディスク装置を複数接続した場合、特定のディスクがアクセスしてしまうと、他のディスクディスク装置が他のデバイスとデータのやり取りを出来なくなる。発明のディスクディスク装置では、ルータを介してディスクを特定し、バスを獲得する方法である。よって特定のチャンネルが障害によりデータ転送が不可能となった場合においても、代用の経路及びディスク装置を見出すことが可能となるため、耐故障特性の向上も可能である。

【0011】

【実施例】請求項1にかかわる一実施例を説明する。図2がブロック構成図であり、図3は図のディスクインタフェース制御部の詳細図である。ここではディスク装置としてハードディスク装置を2台示したが3台以上で使用する場合においても同様に適用が可能である。

【0012】ホストコンピュータHCはディスク装置DKに対する入出力要求を実行する場合、入出力コマンドをコマンドデータエンコード機構1に対して発行する。コマンドデータエンコード機構はホストコンピュータHCから受け取った入出力コマンドをシリアルコマンドバスSCBを介してディスクインタフェース制御部2に対して送出する。入出力コマンドを受け取ったディスクインタフェース制御部2は内部のコマンド解釈部においてそのコマンドがデータ転送を伴うかそうでないかをコマンドコードを元にして判断し、データ転送を伴わない場合はデータ転送経路セレクト22を介してディスク装置DKに必要情報を要求する。得られたデータは再びコマンド解釈部21を経てシリアルコマンドバスSCBを介してコマンドデータエンコード機構1に転送する。

【0013】次に、データの転送を伴うと判断された場合の処理を説明する。まずディスク装置DKより読みだしデータが存在すると判断された場合、データ転送経路セレクト22を介してディスク装置DKに読みだし要求を発行する。同時にコマンド制御部21は転送経路設定機構24に対して、パラレルデータバスPDBの確保の要求と転送するデータの長さの設定を行う。ディスク装置DKよりデータが読みだされてきた時点でコマンド解釈部21はヘッダ付加機構23に対してデータの先頭にディスクインタフェース制御部2及びディスク装置DKの識別コードをデータの先頭部に付加することを要求しそのデータは転送経路設定機構24を介してパラレルデータバスPDB上に流される。ディスク装置DKのデータの読み出しが終了するまで転送経路設定機構はパラレルデータバスPDBを占有し使用するが、この間シリアルコマンドバスSCBはパラレルデータバスの転送状態にかかわらずコマンドの転送が可能となる。コマンド解釈部はデータの転送の終了を待ちディスク装置DKの転送状態を確認した上でコマンドバスSCBを介してホス

5

トコンピュータHCにステータスを転送し一連の動作を終了する。

【0014】次にディスク装置DKにデータを書き込む場合、データ転送経路セクタ22を介してディスク装置DKに書き込み要求を発行する。同時にコマンド制御部21は転送経路設定機構24に対して、パラレルデータバスPDBの確保の要求と転送するデータの長さの設定を行う。一方、コマンドデータエンコード機構1はデータの先頭にディスクインタフェース制御部2及びディスク装置DKの識別コードをデータの先頭部に付加した上でデータを送出する。コマンドデータエンコード機構1が送出したデータはパラレルデータバスPDBを介して転送経路設定機構24に送られる。データが送られてくるとヘッダ付加機構23はデータの先頭に付加されたディスクインタフェース制御部2及びディスク装置DKの識別コードを判定し、コードが異なる場合はその結果をコマンド解釈部に知らせる。コードに問題が無ければデータ転送経路セクタ22を介してディスク装置DKに対してデータの書き込みが実行される。このデータ書き込み動作の間はパラレルデータバスPDBはコマンドデータエンコード機構1により占有され使用されるが、この間シリアルコマンドバスSCBはパラレルデータバスの転送状態にかかわらずコマンドの転送が可能となる。コマンド解釈部はデータの転送の終了を待ちディスク装置DKの転送状態を確認した上でコマンドバスSCBを介してホストコンピュータHCにステータスを転送し一連の動作を終了する。

【0015】次に請求項2にかかわる一実施例を説明する。図4がブロック構成図であり、図5は図4のデータバスルータ3の詳細図である。ここではディスク装置としてハードディスク装置を2台、ホストコンピュータを2台示したが何れも3台以上で使用する場合も同様に適用が可能である。

【0016】ホストコンピュータHCはディスク装置DKに対する入出力要求を実行する場合、入出力コマンドをコマンドデータエンコード機構1に対して発行する。コマンドデータエンコード機構はホストコンピュータHCから受け取った入出力コマンドをシリアルコマンドバスSCBを介してディスクインタフェース制御部2に対して送出する。入出力コマンドを受け取ったディスクインタフェース制御部2は内部のコマンド解釈部においてそのコマンドがデータ転送を伴うかそうでないかをコマンドコードを元にして判断し、データ転送を伴わない場合はデータ転送経路セクタ22を介してディスク装置DKに必要情報を要求する。得られたデータは再びコマンド解釈部21を経てシリアルコマンドバスSCBを介してコマンドデータエンコード機構1に転送する。

【0017】次に、データの転送を伴うと判断された場合の処理を説明する。まずディスク装置DKより読みだすデータが存在すると判断された場合、データ転送経路

6

セクタ22を介してディスク装置DKに読みだし要求を発行する。同時にコマンド制御部21は転送経路設定機構24に対して、パラレルデータバスPDBの確保要求と転送するデータの長さの設定を行う。転送経路設定機構24はデータバスルータ3に対してコマンドデータエンコード機構1の識別番号であるイニシエータ識別コードと、ディスクインタフェース制御部の識別コードであるターゲットコードを渡すとともにデータバスの継続的な確保を要求する。次に、パラレルデータバスPDBの確保方法に関して説明する。データバスルータ3は外部インタフェースアクセス機構33を介して受け取ったバス要求をデータアクセス有無判定機構で受け取り、受け取った外部インタフェースアクセス機構ではないもう一つの外部インタフェースアクセス機構33を介してパラレルデータバスPDBの継続的な確保の要求を発行する。以上の手順により複数のデータバスルータ3を経て最終的に確保されたパラレルデータバスPDBをデュアルポートバッファメモリを介して接続され、最低限のタイムラグでホストコンピュータHCのコマンドデータエンコード機構1とディスクインタフェース制御部2を接続することが可能である。また、これらの処理において、各データバスルータはコマンドデータエンコード機構1とディスクインタフェース機構2の識別コードをモニタすることが可能であるため、図4のように複数のホストコンピュータHCあるいは複数のディスクインタフェース制御部2が混在している場合においてもバスの確保が可能である。

【0018】一方パラレルデータバスPDBの確保が実行されている間に、コマンド解釈部21はディスク装置DKよりデータが読みだされてきた時点でヘッダ付加機構23に対してデータの先頭にディスクインタフェース制御部2及びディスク装置DKの識別コードをデータの先頭部に付加することを要求しそのデータは転送経路設定機構24よりパラレルデータバスPDB及びデータバスルータ3を介してコマンドデータエンコード機構1に送られる。ディスク装置DKのデータの読み出しが終了するまで転送経路設定機構はパラレルデータバスPDBを占有し使用するが、この間シリアルコマンドバスSCBはパラレルデータバスの転送状態にかかわらずコマンドの転送が可能となる。コマンド解釈部はデータの転送の終了を待ちディスク装置DKの転送状態を確認した上でコマンドバスSCBを介してホストコンピュータHCにステータスを転送し一連の動作を終了する。

【0019】次にディスク装置DKにデータを書き込む場合、データ転送経路セクタ22を介してディスク装置DKに書き込み要求を発行する。同時にコマンド制御部21は転送経路設定機構24に対して、パラレルデータバスPDBの確保の要求と転送するデータの長さの設定を行う。転送経路設定機構24はデータバスルータ3に対してコマンドデータエンコード機構1の識別番号で

あるイニシエータ識別コードと、ディスクインタフェース制御部の識別コードであるターゲットコードを渡すとともにデータバスの継続的な確保を要求する。

【0020】次に、パラレルデータバスPDBの確保方法に関して説明する。データバスルータ3は外部インタフェースアクセス機構33を介して受け取ったバス要求をデータアクセス有無判定機構で受け取り、受け取った外部インタフェースアクセス機構33を介してパラレルデータバスPDBの継続的な確保の要求を発行する。以上10
 の手順により複数のデータバスルータ3を経て最終的に確保されたパラレルデータバスPDBをデュアルポートバッファメモリを介して接続され、最低限のタイムラグでホストコンピュータHCのコマンドデータエンコード機構1とディスクインタフェース制御部2を接続することが可能である。また、これらの処理において、各データバスルータはコマンドデータエンコード機構1とディスクインタフェース機構2の識別コードをモニタすることが可能であるため、図4のように複数のホストコンピュータHCあるいは複数のディスクインタフェース制御部2が混在している場合においてもバスの確保が可能である。

【0021】一方、コマンドデータエンコード機構1はデータの先頭にディスクインタフェース制御部2及びディスク装置DKの識別コードをデータの先頭部に付加した上でデータを送出する。コマンドデータエンコード機構1が送出したデータはパラレルデータバスPDB及びデータバスルータ3を介して転送経路設定機構24に送られる。データが送られてくるとヘッダ付加機構23はデータの先頭に付加されたディスクインタフェース制御部2及びディスク装置DKの識別コードを判定し、コードが異なる場合はその結果をコマンド解釈部に知らせる。コードに問題が無ければデータ転送経路セクタ22を介してディスク装置DKに対してデータの書き込みが実行される。このデータ書き込み動作の間はパラレルデータバスPDBはコマンドデータエンコード機構1により占有され使用されるが、この間シリアルコマンドバスSCBはパラレルデータバスの転送状態にかかわらずコマンドの転送が可能となる。コマンド解釈部はデータの転送の終了を待ちディスク装置DKの転送状態を確認した上でコマンドバスSCBを介してホストコンピュータHCにステータスを転送し一連の動作を終了する。

【0022】次に請求項3にかかわる1実施例を説明する。図6が本発明に使用するデータのフォーマットであり、これらはコマンドバス上で相互のメッセージのやり取りを行う際に使用するフォーマットである。図7は本発明にかかわるデータ処理フローである。図4が使用するハードウェアのブロック構成図である。ここではディスク装置としてハードディスク装置を2台、ホストコンピュータを2台示したが各々3台以上で使用する場合に

においても同様に適用が可能である。

【0023】図7はホストコンピュータHCとディスク制御インタフェース部2と、データバスルータ3との間でのやり取りする処理プロトコルを表わす。またホストコンピュータコマンドデータエンコード機構1と、ディスクインタフェース制御機構2と、データバスルータ3との間でのやり取りは、CIFコマンドシリアルインタフェース上でやり取りを行う。コマンドフォーマットとしては、図2のコマンドフォーマットのようにディスクのI/O制御コマンドとデータ転送経路制御コマンドからなるが、データの転送フェーズを伴わない場合はデータ転送経路制御コマンドは不要である。また、データ転送フェーズを伴う場合は、データバスルータの転送経路の確保処理をデータ転送制御コマンドを元にして実行し、転送を行う。

【0024】ここでは、データ転送を伴う場合のコマンド実行シーケンスについて説明する。ホストコンピュータHCからのデータ入出力要求コマンドは、コマンド・データエンコード機構1内でディスクI/O制御コマンドとデータ転送経路制御コマンドを図6のコマンドフォーマットのようにパケット化しシリアルコマンドバスSCB上に流す。ここで、コマンドの転送はシリアルインタフェースのプロトコルには特に依存しない。コマンドパケットは、シリアルコマンドバスSCBを介して、入出力対象であるディスクインタフェース制御部2の識別コードを持つデバイスにたいして、コマンドパケットを発行し、データ転送要求の受付確認コードが返送されるのを待つ。コマンドパケットを受けたディスクインタフェース制御部2は、コマンド要求に基づき、ディスク装置の読みだし/書き込みの準備を行い、転送経路の確保が出来た時点で、ホストコマンドデータエンコード機構に対して、要求受付確認コードを発行する。転送経路の確保シーケンスに関しては後述する。要求受付確認コードを受けたコマンドデータエンコード機構1は転送を行う経路を確認し、データ転送要求をディスクインタフェースに対して発行する。ディスクインタフェースは、データパケットに対して転送コードを付加しデータバス上のコマンドデータエンコード機構1あるいは転送経路の確保に従い、データバスルータ3に対してデータ転送を実行する。コマンドデータエンコード機構1はデータバス上を送られてくる転送コードを確認し、その転送コードがコマンドパケット発行時に取得した転送コードと同じものであるかをモニターし、もし同じ場合はそのデータパケットをコマンドデータエンコード機構1内に取り込む。データの転送終了後、ディスクインタフェース側から送られるステータスを確認し転送フェーズを終了する。

【0025】次にデータ転送の実行の際にディスクインタフェースが行う転送経路確保処理手順について説明する。ディスクインタフェース制御部2の制御処理シーケ

ンスにおいて転送経路を確保する場合は、直接イニシエータデバイスであるコマンドデータエンコード機構1を検索し見つからない場合は、データバスルータ3に対して、別のデータ転送経路の検索を要求できる。よって、複数のデータバスルータ3経由でデータ転送経路の確保が実行される。各データバスルータ3は転送経路の確保を要求されると、イニシエータデバイスを検索する。もし見つからなかった場合は、更にデータバスルータ3を検索し、その間のデータ転送バスを予約した後に、更に他のルータに対して転送経路の確保を要求する。イニシエータデバイスが見つかった場合は、データバスルータ3は自身のコードをルーティングリストに付加し、ルーティングリスト内に不要な転送経路があった場合はその箇所を除いた上で、転送経路確保を要求してきたデータバスルータ3あるいは、ディスクインタフェース制御部2にルーティングリストを返す。このシーケンスにより、最終的にターゲットデバイスに戻されるルーティングリストに基づいてディスクインタフェース制御部2は転送経路をイニシエータであるコマンドデータエンコード機構1に対して送出する。

【0026】次に請求項4にかかわる1実施例を説明する。図8がブロック構成図であり、図9は図8のディスクインタフェース制御部4の詳細図である。ここでは、ディスク装置は3台、ホストコンピュータは3示しているが、何れも3台以上接続する場合も同様に構成可能である。

【0027】ホストコンピュータHCはディスク装置DKに対する入出力要求を実行する場合、入出力コマンドをディスクインタフェース制御部4に対して発行する。入出力コマンドを受け取ったディスクインタフェース制御部4はホストインタフェース制御機構43においてそのコマンドがデータ転送を伴うかそうでないかをコマンドコードを元にして判断し、データ転送を伴わない場合は直接ディスク装置DKに対して情報を要求する。その際にディスクインタフェース制御部2に直接接続されたディスク装置DKに対する要求の場合はディスクインタフェース45を介してディスク装置DKに必要情報を要求し、得られたデータはホストインタフェース制御機構43を経てホストコンピュータHCに転送される。直接接続されていないディスク装置DKの場合には外部コマンドバス6を介して接続された隣接したディスクインタフェース制御部4に対してディスクの情報を要求する。このような隣接したディスクインタフェース制御部4に対する外部からのアクセスは、データ有無判定機構41を介して全て判断がなされる。その際、外部からのアクセス要求が直接接続されたディスク装置DKに対する要求かどうかを判断を行い、そうであった場合にはディスクインタフェース45を介してディスクの情報を要求し、そうでない場合には更に隣接するディスクインタフェース制御機構に対して外部コマンドバスを使用し、デ

ィスクの情報を要求する。このように、隣接したディスクインタフェース制御部でリレー式に情報を要求できる。

【0028】次に、データの書き込み動作に関して説明する。直接接続されたディスク装置DKが書き込み対象のディスク装置と判断された場合は、ディスク装置DKに書き込み要求を発行する。書き込みデータは直接ホストインタフェース制御機構43からディスクインタフェース46に転送される。一方、直接接続されたディスク装置DKが書き込み対象のディスク装置と判断されなかった場合には、隣接したディスクに対してデータが存在するかを外部コマンドバスを介して問い合わせる。問い合わせを受けたディスクインタフェース制御機構4内のデータ有無判定機構41は、判断結果を問い合わせたディスクインタフェース制御機構に返すとともに、書き込み対象のディスク装置があると判断された場合には更に隣接したディスクインタフェース制御部4に対してデータの有無を問い合わせる。その結果を受けた発行元のディスクインタフェース制御部4は、データが存在するディスクインタフェース制御部4へのデータの書き込みを要求するとともに、ホストインタフェースからデュアルポートバッファメモリ42に対してデータの書き込みを実行する。この間内部インタフェースデータ転送バス48は継続的にデータ転送に割り付けられる。同時に外部インタフェースアクセス機構B44に対して、デュアルポートバッファメモリのデータを読みだし隣接するディスクインタフェース制御部4に対して転送を実行する。その間、外部インタフェースデータ転送バスは継続的に外部インタフェースアクセス機構及びデュアルポートメモリ42に対して割り当てられる。また、直接該当ディスク装置の存在しないディスクインタフェース制御部4ではデータ有無判定機構により外部インタフェースアクセス機構A44及び外部インタフェースアクセス機構B45に対して外部インタフェースデータ転送バス47によりデータの転送を行うよう指示される。この間外部インタフェースデータ転送バス47は継続的に外部インタフェースアクセス機構間のデータの転送に割り付けられる。書き込み対象のディスク装置が接続されているディスクインタフェース制御部4では、外部インタフェースアクセス機構A44から読みだしたデータをいったんデュアルポートメモリ42に書き込み、そのデータをディスクインタフェース46がディスク装置DKに転送を実行する。その動作結果は外部コマンドバスを経て入出力要求の発行元ディスクインタフェース制御部4に転送され一連のデータ入出力を終了する。

【0029】

【発明の効果】本発明は、コンピュータのディスク装置に上位レベルのデータ入出力プロトコル方法を実現することにより、記録データの長さにかかわらずレスポンスのよりディスク装置のインタフェースを実現する。この

11

ような機能をディスク装置に持たせることにより、従来ホストコンピュータおよびディスク装置では性能低下の原因となっていたI/Oバスのボトルネックの解消を可能とする。また、ホストコンピュータからは、転送バス経路のなどを管理する必要が無いため、ハードウェアの構成に依存しないドライバソフトウェアの開発が可能となる。

【図面の簡単な説明】

【図1】従来技術の説明にかかわるディスク装置のデータ転送方法の例を説明するための構成図である。

【図2】請求項2にかかわる、自動ルーティングディスク装置の一実施例を説明するためのブロック図である。

【図3】図2のディスクインタフェース制御部の詳細図である。

【図4】請求項3にかかわる、自動ルーティングディスク装置の一実施例を説明するためのブロック図である。

【図5】図4のデータバスルータの詳細図である。

【図6】請求項1にかかわる、データ転送シーケンスを実行する際に使用するコマンドパケットのフォーマットを示す。

【図7】請求項1にかかわる、データ転送の処理シーケンスを示すブロック図である。

【図8】請求項4にかかわる、自動ルーティングディスク装置の一実施例を説明するためのブロック図である。

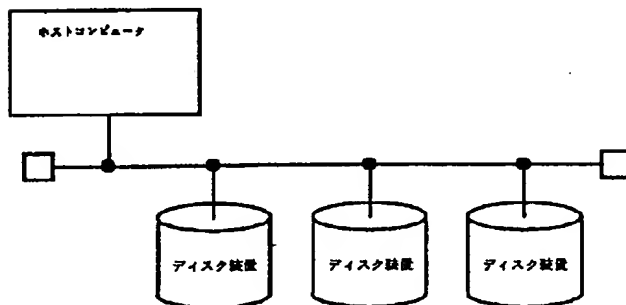
【図9】図8におけるディスクインタフェース制御部の詳細図である。

【符号の説明】

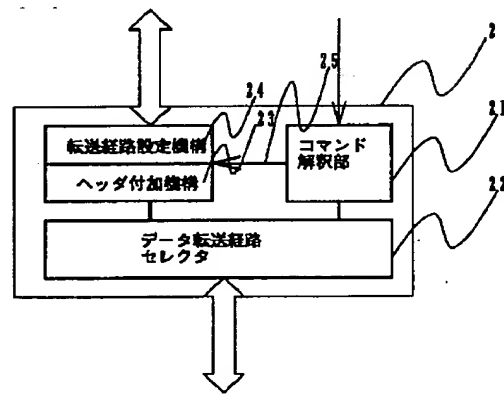
12

- 1 コマンドデータエンコード機構
- 2 ディスクインタフェース制御部
- 21 コマンド解釈部
- 22 データ転送経路セクタ
- 23 ヘッダ付加機構
- 24 転送経路設定機構
- 3 データバスルータ
- 31 データ有無判定機構
- 32 デュアルポートバッファメモリ
- 33 外部バスインタフェースアクセス機構
- 34 外部インタフェース制御信号
- 35 データ転送バス
- 4 ディスクインタフェース制御部
- 41 データ有無判定機構
- 42 デュアルポートバッファメモリ
- 43 ホストインタフェース制御機構
- 44 外部インタフェースアクセス機構
- 45 外部インタフェースアクセス機構
- 46 ディスクインタフェース
- 47 外部インタフェースデータ転送バス
- 48 内部インタフェースデータ転送バス
- 5 外部データバス
- 6 外部コマンドバス
- DK ディスク装置
- PDB パラレルデータバス
- SCB シリアルコマンドバス
- HC ホストコンピュータ

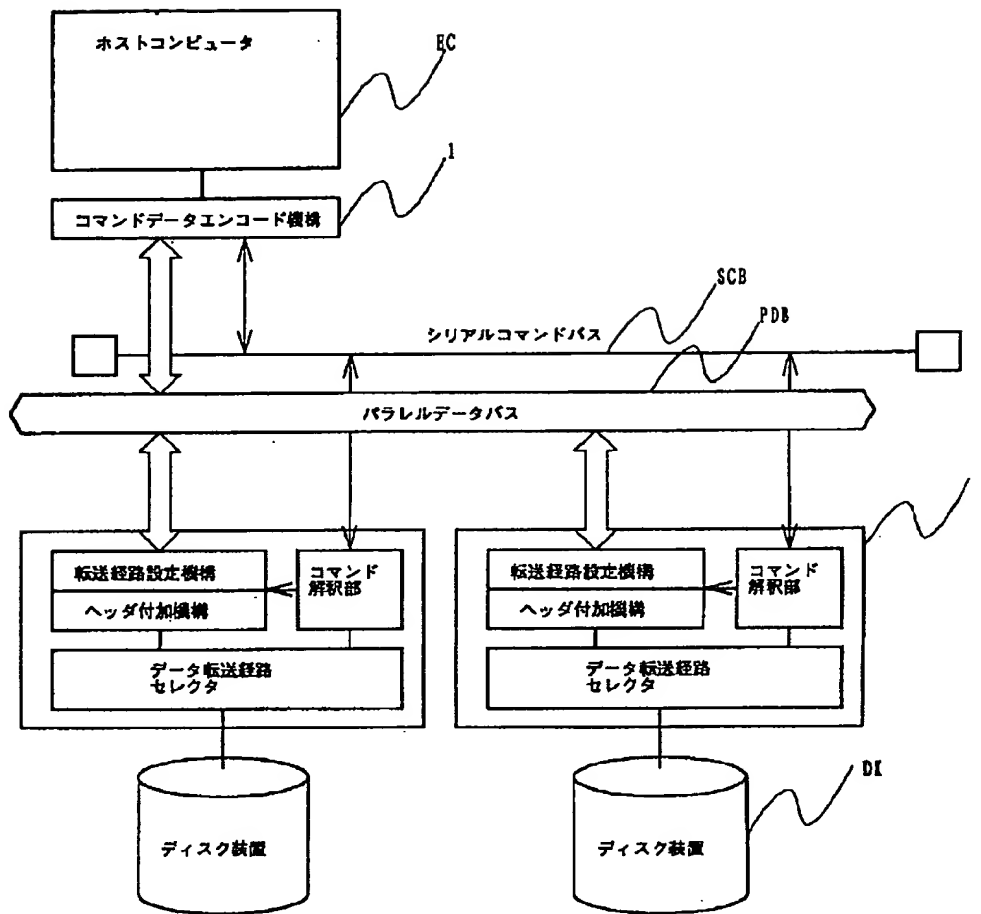
【図1】



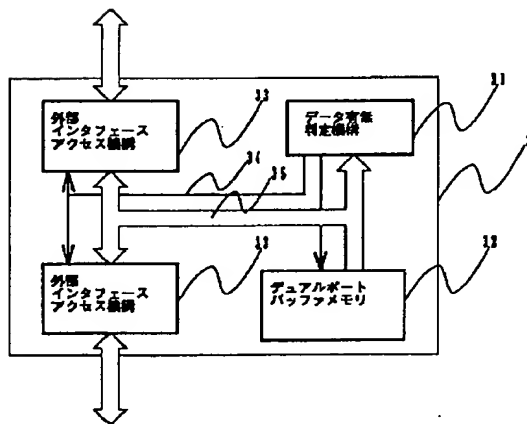
【図3】



【図2】



【図5】



【図6】

コマンドパケットフォーマット

コマンド 表	コマンド 識別フラグ	ディスクI/O制御コマンド	データ転送経路制御コマンド 又は 転送要求受け付けコマンド
-----------	---------------	---------------	----------------------------------

ディスクI/O制御コマンド

コマンド コード	(ブロックアドレス)	(ブロック長さ)	ステータス
-------------	------------	----------	-------

データ転送経路制御コマンド

イニシエータ コード	ターゲット コード	メトリックス
---------------	--------------	--------

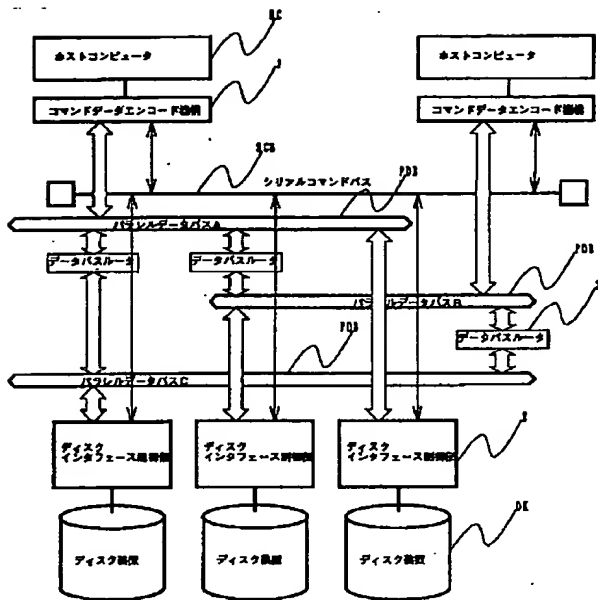
ルーティンダリスト付データ転送制御コマンド

イニシエータ コード	ターゲット コード	メトリックス	データバス ルートコード	...	データバス ルートコード
---------------	--------------	--------	-----------------	-----	-----------------

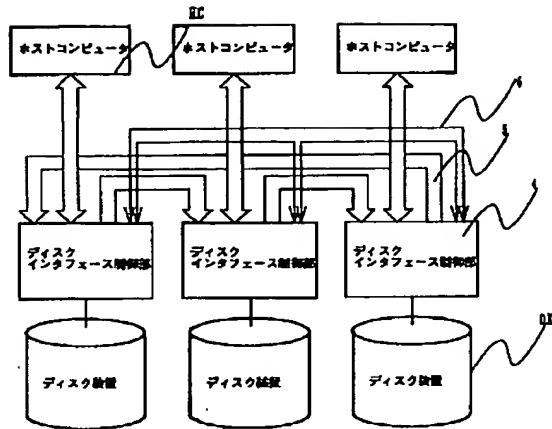
転送要求受け付けコマンド

データ転送 コード	データ転送 モード	バースト転送長
--------------	--------------	---------

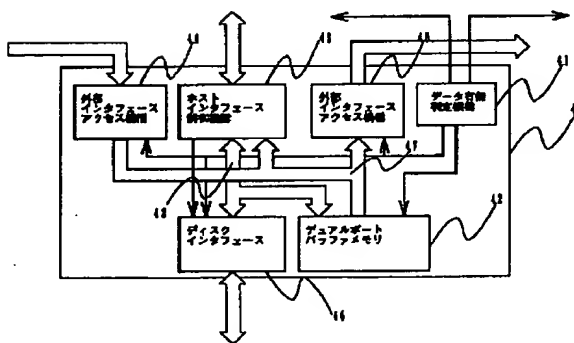
【図4】



【図8】



【図9】



【図7】

